



**OnSSI** On-Net Surveillance Systems Inc.

# MJPEG vs MPEG4

**Understanding the differences,  
advantages and disadvantages of each  
compression technique**

**White Paper**

On-Net Surveillance Systems Inc.  
222 Route 59 Suite 303, Suffern, NY 10901  
Phone: (845) 369-6400 | Fax: (845) 369-8711  
[www.onssi.com](http://www.onssi.com)

## MJPEG vs MPEG4

### Overview

This white paper is intended to help the reader understand the roles and benefit of both the MJPEG and the MPEG-4 compression methods, and when to use which one based on necessity and performance. It begins with some basic concepts that relate to both image size (in pixels) and PPF (pixels per foot).

This is followed by some background information on compression: why it was needed, and how it is accomplished; finally, each compression method is described separately, listing the key advantages of each.

### Understanding image size and dimension - CIF

The maximum number of pixels that make up the image is determined by the camera's CCD (Charge Coupled Device) - an integrated circuit containing hundreds of thousands, or millions, of pixels that convert light energy into electronic signals.

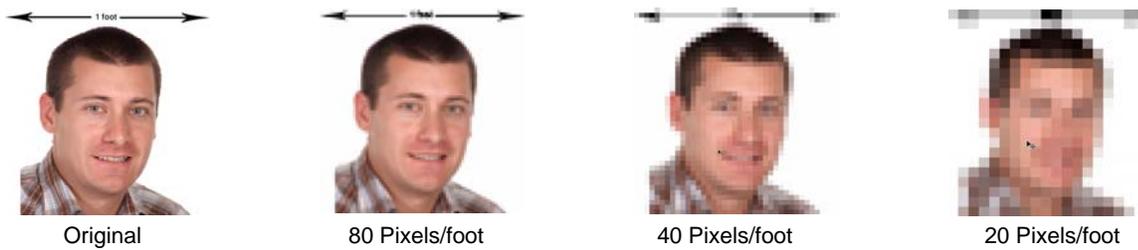
### CIF – (Common Intermediate Format or Common Image Format)

- CIF, the NTSC standard for broadcast TV, contains 352x240 (=84,480) pixels.
- The original CIF is also known as Full CIF (FCIF). Image resolutions can be fractions of 1CIF (e.g. Quarter CIF, QCIF) and multipliers of 1CIF (2CIF.... 16CIF).
- Frame Size: When using 1 CIF, an image size of 8kb to 10kb is the optimal range. When using 4CIF, anything between 30kb and 40kb is optimal for 98% of all installations. These frame sizes have been industry tested and any increase in size only adds enhancements to the color depth of the images, which is basically undetectable with the human eye for video surveillance implementation. You will only impact HDD storage and CPU resources.

### Pixels Per Foot (PPF)

One major purpose of video surveillance is in essence coming up with images that provide enough detail to identify the person or object captured. The level of detail is defined in terms of Pixels Per Foot (PPF). For example, a 'headshot' image (see example) of about 1x1 feet, at 40 PPF, will consist of 1600 (40 x 40) pixels. Rules of thumb:

- Facial recognition requires at least 40 PPF.
- License plate recognition: at least 60 PPF.



Once you've determined the image detail required (in PPF), you need to determine the size of the coverage area. For instance, a requirement may be that facial recognition (40 PPF) would be enabled within a particular area (coverage frame size) 9 feet wide. The relative area height would be about 6.1 feet high (based on the NTSC width/height ratio).

The image size calculation (number of pixels per image/frame) would be:

40 PPF = 1600 pixels/sq. ft.

Viewed area = 9 x 6.1 = 55 sq. ft.

Required resolution: 55 sq. ft. x 1600 pixels/sq. ft. = 88,300 pixels

### **Image Compression - Background**

An ordinary, uncompressed analog video stream (when digitized according to the CCIR standard 601) can consume as much as 165 Mbps of bandwidth (megabits per second). Video compression techniques have been developed to reduce this high bandwidth.

A full size (4CIF), full rate video stream (30 FPS), using average compression (32KB frame size) is about 8 Mbps, which is significantly less than its uncompressed size. For most security applications, the rate can be cut to a quarter of that (7 1/2 FPS), to about 2 Mbps.

Increasing the compression level causes degradation of the image. This would be manifested in visual inaccuracies, referred to as digital artifacts. A balance needs to be drawn between reducing image size and keeping the image clear. Typically, the defaults used by network camera and video server manufacturers have achieved this balance, and standards have been established for manufactures to follow.

The foundation of these standards is JPEG, named after the group that developed it in the mid-1980s, the Joint Photographic Experts Group. The first JPEG standard was released in 1991. Another group of developers, in the late 1980s, developed a standard for compressing moving pictures and audio - the Motion Picture Experts Group, hence the name MPEG. MPEG-4 was released in 1998.

**Video compression is accomplished by** removing color nuances within the image, reducing the color resolution, removing invisible parts of the picture, and in the case of MPEG standards, using video sequencing compression—transporting only the changes in the sequence.

For the most part, it is the video sequencing compression that allows MPEG compression to use less network bandwidth and storage than MJPEG.

### **MOTION JPEG (M-JPEG)**

Motion JPEG (M-JPEG) is a video codec where each video field (frame) is separately compressed into a JPEG image. The resulting quality of video compression is independent from the motion in the image (which differs from MPEG video where quality often decreases when footage contains lots of movement).

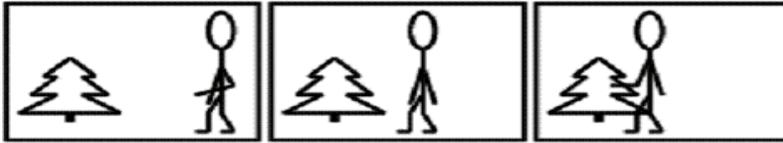
Characteristics of M-JPEG):

- At low bandwidth availability, priority is given to image resolution (i.e. transmitted images would maintain their quality, however some images would be dropped).
- Minimum latency in image processing.
- Images have a consistent file size.
- Still the most widespread picture compression format used today

Similar to a digital still picture camera, a network camera captures individual images and compresses them into a JPEG format. The camera can capture and compress, for example, 30 such individual images per second (30 fps), and then make them available as a continuous flow of images over a network to a viewing station.

At a frame rate of about 16 fps and above, the viewer will perceive full motion video. As each individual image is a complete JPEG compressed image, they will all have the same guaranteed quality, determined by the compression level as defined for the network camera or network video server.

**Example of a sequence of three complete JPEG images**

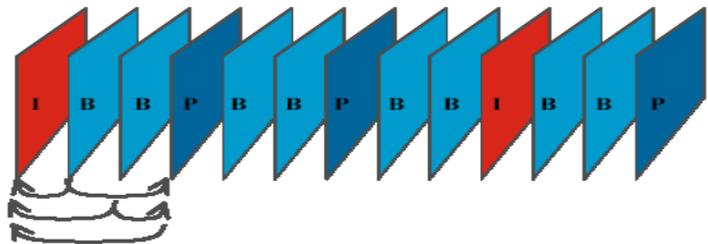


**MPEG-4**

MPEG-4, introduced in late 1998, is the designation for a group of audio and video coding standards and related technology agreed upon by the ISO/IEC. The primary uses for the MPEG-4 standard are web (streaming media) and CD distribution, and broadcast television. Key points of MPEG-4 are the following:

- Uses key-frames and object oriented differential compression format
- At low bandwidth availability, priority is given to frame rate
- Bandwidth is directly affected by amount of motion in the frame (since frames may differ in size)
- May raise problems when used as evidence in court, since non-keyframe frames contain only the differential between the previous and subsequent frames.
- Some proprietary implementations and licensing is required.

Simply described, MPEG's basic principle is to compare compressed images to be transmitted over the network, using the first compressed image as a reference frame (called an I frame), sending it and only the parts of following images that differ from the reference image (B and P frames).



The network viewing station will then reconstruct all images based on the reference image and the “difference data” contained in the B and P frames. A typical sequence of I, B, and P frames may look as below. Note that a P frame may only reference a foregoing I or P frame, while a B frame may reference both foregoing and coming I and P frames:

A result of applying MPEG video compression is that the amount of data transmitted across the network is less than that of Motion JPEG. This is illustrated below where only information about the differences in the second and third frames is transmitted.

MPEG is in fact far more complex than indicated above, and parameters such as prediction of motion in a scene and identifying objects are additional techniques or tools used within MPEG.



### **Key Advantages of MPEG-4**

- Reduces the amount of storage needed
- Increases the amount of time video can be stored
- Reduces the network bandwidth used by the surveillance system

### **Examples of MPEG-4 applications**

An example of a situation where MPEG-4 could be utilized to solve a bandwidth issue is if the video traffic was going to be part of an existing 100Mbs network. The network would be saturated sooner if MJPEG cameras were put on it.

Another example of where MPEG-4 could be the solution is if the storage on the NVR was relatively small and archiving could only be done once a day.

A third example of utilizing MPEG-4 benefits would be if the length of time desired to store archives was unusually long, or just too long, if MJPEG images were stored.

### **Key Advantages of MJPEG**

- Ideal for courtroom evidence
- Frame by frame playback offers more frames to view
- Technology is simpler; this can lower the price of a camera or video server
- Less delay issues when used with audio
- At low bandwidth availability, priority is given to the Image Resolution
- Easier to work with if using photo editing s/w

### **Examples of MJPEG applications**

An example for utilizing MJPEG complete frame by frame playback benefit would be if the video playback was going to be used in court cases; MPEG-4 could be challenged.

Another example of when to use MJPEG is if staying within a budget for a project means buying cameras without MPEG-4 capability, but having other key features desired.

A third example is if in a situation where Image Resolution is the priority and the bandwidth is limited, then MJPEG would be the way to go.

### **Terminology and Acronyms**

**Interlaced** – A technique used in old television system where the picture is divided into two half pictures containing every other line each. When displayed, first the odd lines are displayed then the even lines followed by the odd lines of the next picture and so on.

**Progressive Scan** – Each picture in the video sequence is the full picture displayed all in once. This is the opposite of *Interlaced*.

**PAL** – Phase Alternating Line. This is the standard for the analog television format used in Europe with 625 lines at 50 half-pictures per second, i.e. *Interlaced* video.

**NTSC** – National Television Standards Committee. This is the standard for the analog television format used in the US with 525 lines at near 60 pictures per second

**CCIR 601** – A standard for digital video for picture size of 720 × 485 at 60 interlaced pictures per second or 720 × 576 at 50 interlaced pictures per second.

**CIF** – Common Intermediate Format. Video of picture size 352 × 288

**MPEG** – Motion Picture Experts Group. The committee responsible for developing the MPEG standards. Homepage at: [www.mpeg.telecomitalia.com](http://www.mpeg.telecomitalia.com)

**JPEG** – Joint Photographic Experts Group. The committee responsible for developing the JPEG and JPEG 2000 standards. Homepage at: [www.jpeg.org](http://www.jpeg.org)

**ISO** – International Standards Organization. A worldwide federation of national standards bodies from some 140 countries. Homepage at: [www.iso.ch](http://www.iso.ch)

**IEC** – International Electro-technical Commission. International Electro-technical Commission. An international standards and conformity assessment body for all fields of electro technology. Homepage at: [www.iec.ch](http://www.iec.ch)

**ITU** – International Telecommunications Union. An international organization within the United Nations System where governments and the private sector coordinate global telecom networks and services. Homepage at: [www.itu.int](http://www.itu.int)

© 2002-2006, On-Net Surveillance Systems Inc. The information contained in this document is provided as general information only. On-Net Surveillance Systems Inc. makes no claims, promises or guarantees about the accuracy, completeness, or adequacy of the contents of this documents and expressly disclaims liability for errors and omissions in its contents.